

Protein with cell proliferation and cell division modulating activity and DNA encoding such protein

Specification

The present invention relates to DNA sequences, expression vectors containing such DNA sequences, proteins encoded thereby, the use of these proteins for inducing oocyte maturation or modulating cell division and in a pharmaceutical composition, uses as diagnostic markers or for identifying substances modulating the cell cycle progression and/or cell proliferation and/or differentiation, as well as further applications derived therefrom.

Proteins influencing cell division, proliferation or differentiation are generally of great interest. These substances open up a variety of possible uses which can be of interest for several applications depending on the specificity of these proteins. Usually, drugs which make use of the effect of suitable proteins to control or prevent pathological situations can also be derived therefrom. In general, newly found proteins which can be produced recombinantly are therefore received with great interest. They do not only have potential pharmaceutical effects themselves but can often also be used as diagnostic means or as means for developing secondary pharmaceutical agents.

With the current systematical elucidation of the sequences of the human genome (human genome project), many sequences are found which obviously code for proteins. In most cases, though, the function of these proteins is completely unknown, so that it cannot be foreseen which possible uses such products might have.

- 2 -

The object underlying the present invention was to detect sequences coding for proteins influencing the cell cycle, cell division and cell proliferation. It was a further object of the present invention to produce corresponding proteins.

5

In accordance with the present invention this object was accomplished by providing a DNA sequence characterized in that it contains:

10

- (a) a sequence as shown in SEQ ID NO.1 or 2,
- (b) a sequence which encodes the same protein as (a) but is degenerate as a result of the genetic code,
- (c) a sequence hybridizing under stringent conditions to the sequences of (a) and/or (b),
- (d) a genomic sequence containing the sequence of (a), (b) or (c) and further containing one or more introns,
- 15 (e) a sequence which differs from (a), (b), (c) or (d) due to its origin from a different species.

20

25

30

In order to find proteins involved in cell cycle activation, a *Xenopus* oocyte cDNA was prepared and cloned in expression vectors. The primary library was subdivided into pools and plasmid DNA was purified from the pools and in vitro transcribed to obtain mRNAs. The mRNA pools were injected into stage VI oocytes which were incubated to allow for protein expression. Pools which upon microinjection in oocytes were capable of inducing oocyte maturation on their own or of strongly accelerating progesterone-induced maturation were subdivided into smaller pools and reinjected until single positive clones were isolated. Following this approach, out of a huge number of mRNA pools two specific sequences corresponding to SEQ ID NOs.1 and 2 were isolated. These sequences do not correspond by DNA hybridization experiments to any known proteins inducing oocyte maturation, including protein kinase Mos, the protein phosphatase cdc25 and several A and B type cyclins. The mRNA prepared from the two isolated clones containing SEQ ID NO.1 or 2 was used for protein expression and

- 3 -

the obtained proteins were found capable of potently inducing oocyte maturation also in the absence of progesterone stimulation. The obtained DNA sequence data as shown in SEQ ID NOs.1 and 2 demonstrate that these two clones contain open reading frames coding for related proteins.

5

The present invention comprises the sequences shown in SEQ ID NOs.1 and 2 which, however, may also contain certain deviations. In particular, the present invention covers deviations which are present in the DNA only but which, owing to the diversity of the genetic code, encode the same protein as SEQ ID NO.1 or 2. Furthermore, the present invention comprises sequences which hybridize under stringent conditions with SEQ ID NO.1 or 2, or sequences deviating therefrom as set out under (b). The present invention also comprehends the corresponding genomic sequences of the cDNA sequences of SEQ ID NO.1 or 2, or of sequences deviating therefrom as set out under (b) or (c). Such genomic sequences may contain one or more introns which are cleaved off during translation and processing and thus do not influence the finally encoded protein.

10

15

20

Still further, the present invention comprises sequences which deviate from those of (a), (b), (c) or (d) owing to their origin from a different species.

25

Very often, highly conserved DNAs coding for proteins which have the same activities in different species, such as mouse or human, show only slight differences. In most cases, deviations occur only in some nucleotides and/or in few amino acids of the coded protein. Hence, by means of the concretely disclosed sequences as of SEQ ID NO.1 or 2, corresponding nucleic acids in other species, which code for proteins with the same or a very similar activity, can easily be found. Such similar sequences are therefore comprised by the present invention, too.

30

- 4 -

The DNA sequences according to the invention encode proteins capable of inducing oocyte maturation and/or modulating cell division, proliferation and/or differentiation.

5 The term "modulating" according to the invention is meant to encompass promotion as well as inhibition. Experimental results show that proteins encoded by the DNA according to the invention show in some instances a promoting, in other instances an inhibiting effect on cell division.

10 In a preferred embodiment of the invention the DNA sequences further contain expression controlled sequences which are operably linked to the coding DNA sequence. Any suitable expression control sequences may be used for the present invention. Particularly preferred sequences are those which allow a favourable control of expression, such as sequences allowing
15 induction of expression or inhibition of expression. Induction or inhibition generally takes place via the binding of a respective inductor or inhibitor molecule to operator sequences. Corresponding expression control sequences are known to the person skilled in the art, the lac operator being an example therefor.

20

A further subject matter of the present invention is an expression vector containing a DNA sequence according to the invention.

25

As set out above for the DNA sequences, the expression vector also particularly preferably comprises expression control sequences allowing for specific expression control. Also, sequences that allow for positive selection of transformed host cells are known to the man in the art and are preferably introduced in the expression vectors according to the invention.

30

A further subject matter of the present invention is a protein encoded by a DNA sequence according to the present invention. As explained above, the protein according to the invention, which preferably contains an amino acid

- 5 -

sequence according to SEQ ID NO.3 or 4, modulates cell proliferation and differentiation. Oocyte maturation is induced by the proteins of the invention whereas expression in some mammalian cell lines seems to inhibit cell division. The particularly preferred proteins of SEQ ID NOs.3 and 4 are capable of inducing oocyte maturation considerably faster than the same amount of injected *malE-Mos* or progesterone treatment. The entire cell cycle in *Xenopus* oocytes is extraordinarily strongly activated by the proteins of the invention. Only low amounts of the protein of the invention are required to stimulate oocyte maturation.

The protein according to the present invention may have deletions, substitutions and/or additions of amino acids in regions which do not affect the activity. However, the activity of the protein must not be considerably impaired thereby. Further, it is preferred that at the most up to 5 % of the amino acid content of the protein of the invention has deletions, substitutions and/or additions of amino acids. It is not difficult for the skilled artisan to find out which regions may contain deletions, substitutions or additions. Corresponding changes can be made in the nucleic acids, followed by expression and an activity test. By means of site-directed mutagenesis manifold variants can easily be produced and expressed. The person of skill in the art can easily simultaneously test a multitude of such mutants for their activity (high-throughput screening), whereby as a prerequisite at least half the activity of the proteins shown in SEQ ID NO.3 or 4 has to be retained. By means of computer-aided conformation studies the regions of the protein which are less probably involved in the activity of the protein can be determined. Particularly in such regions can mutations be made.

A still further subject matter of the present invention is the use of the protein according to the invention for modulating oocyte maturation and/or promoting cell division, cell proliferation or cell differentiation.

- 6 -

As a further subject matter of the present invention a pharmaceutical composition may be formulated on the basis of this possible use. The pharmaceutical composition according to the invention contains as active agent a protein according to the invention which in particular contains the amino acids as of SEQ ID NO.3 or 4 or sequences derived therefrom which may exhibit the aforementioned mutations, deletions or substitutions.

Preferably, the pharmaceutical composition contains the protein in combination with a pharmaceutically acceptable carrier or adjuvant.

The pharmaceutical composition according to the invention may be used for all pathological situations in which it is desired to modulate cell proliferation, cell differentiation or cell maturation. Examples for such applications are the promotion of growth and maturation of specific cell types, e.g. ovarian cells, so that the pharmaceutical composition of the invention is in particular also useful and suitable for fertility treatments.

The protein according to the invention can further be used as diagnostic marker for cell proliferation and/or cell differentiation. The amount of said protein contained in an organism can be correlated to the cell proliferation or differentiation rate. As soon as a basic value has been determined, the amount of this protein present in, e.g., different development stages of cells can be determined, thus showing the particular development status.

A further possible use of the proteins according to the invention lies in their capability of acting as a target for the identification of drugs modulating cell cycle progression and/or cell proliferation and/or cell differentiation. By means of these proteins a multitude of substances can be tested for their modulation capability. To this end, a system is provided which comprises cells susceptible to the proteins of the invention, proteins according to the invention and a substance which is to be examined as to its modulating activity. It can then be determined whether the activity of the protein to

- 7 -

modulate cell proliferation and/or differentiation is weakened or even prevented by said substance. Using high-throughput screening (HTPS), such experiments can be carried out for a multitude of substances simultaneously. A particularly preferable use for identifying substances modulating cell proliferation and/or differentiation lies in the development of drugs for the treatment of cancer or other pathological situations with uncontrolled cell proliferation.

Especially carcinoma grow by uncontrolled cell division, and the inhibition thereof is highly desired. Substances allowing to block such division can be found using the protein of the invention. Even the protein itself might be applicable for inhibiting uncontrolled cell growth in cancers.

Another subject matter of the present invention is the use of the DNA sequences according to the invention as diagnostic marker for the cell proliferation and/or cell differentiation status, whereby the amount of homologous nucleic acids present in the cell is determined by hybridization experiments. Of particular interest is the amount of mRNAs hybridizing to the DNA according to the invention. For this purpose, preferably the DNA sequence according to the invention or a part thereof is labelled, so that after performance of the hybridization experiment the formed double strands may be easily detected. Particularly preferably, a single-stranded DNA sequence should be used corresponding to the antisense strand of the DNA according to the invention. Using such an antisense strand DNA which is complementary to the mRNA, the actual amount of formed protein can be determined on a nucleic acid basis.

In combination with the figures the following examples are to further illustrate the present invention.

Fig. 1 shows a sequence comparison of proteins Is26 and Is27.

- 8 -

Fig. 2 shows the result of experiments with injection of small amounts of recombinant malE-ls26 protein (10 ng) into *Xenopus* oocytes, leading to oocyte maturation considerably faster than the same amount of injected malE-Mos or a progesterone treatment. Fig. 2 also shows the same experiment where cycloheximide preincubation blocked malE-Mos-induced GVBD (germinal vesicle breakdown) but had no effect on ls26-triggered oocyte maturation.

Fig. 3 shows the malE-ls26-induced activation of MAP kinase and cdc2/cyclin B in an immunoblot and by direct measurement of the in vitro kinase activity using MBP and histone H1 as substrates for MAP kinase and cdc2/cyclin.

Fig. 4 also shows an immunoblot using anti-cdc2 antibodies and an in vitro histone H1 kinase assay.

Fig. 5 shows a pull-down experiment using rabbit reticulocyte lysates which demonstrates that ls26 can directly bind to B-type cyclins.

Example 1

To identify novel proteins implicated in cell cycle activation in *Xenopus* oocytes, an expression cloning strategy was used where a *Xenopus* oocyte cDNA library was constructed in the FTX5 expression vector. The primary library was subdivided into pools of 150-200 colonies and plasmid DNA was purified from the pools and in vitro transcribed to obtain mRNAs. The mRNA pools were injected into stage VI oocytes which were incubated for 30 - 36 hours to allow protein expression from the injected mRNAs prior to stimulation with progesterone. Finally, those pools which upon microinjection in oocytes were capable either of inducing oocyte maturation

- 9 -

on their own or of strongly accelerating progesterone-induced maturation were subdivided into smaller pools and reinjected until single positive clones were isolated.

5 Using this approach, out of 105 mRNA pools injected in oocytes two clones were isolated which did not correspond by DNA hybridization experiments to proteins that are known to induce oocyte maturation including the protein kinase Mos, the protein phosphatase cdc25 and several A and B type cyclins. The mRNAs prepared from the two isolated clones, which were
10 referred to as ls26 and ls27, were capable of potentially inducing oocyte maturation in the absence of progesterone stimulation. DNA sequencing showed that these two clones contained open reading frames that encode for related proteins and were fused in frame to the C-terminus of the myc tag in the FTX5 vector.

15 Full-length ls26 and ls27 cDNAs were cloned from a λ ZAP Xenopus oocyte cDNA library using as probes the two cDNAs isolated in the expression screening. The ls26 clone was 1574 base pairs and encoded for a protein of 300 amino acids (SEQ ID NO.1), whilst ls27 was 1357 base pairs in
20 length and encoded for a protein of 298 amino acids (SEQ ID NO.2). Both clones contained stop codons upstream of the first ATG and in the same frame (underlined in SEQ ID NO.1 and 2). The predicted ls26 and ls27 proteins were 91% identical (Fig. 1). When the ls26 and ls27 sequences were tested against DNA and protein sequence data bases, no significant
25 homologies (Blast search) could be detected, suggesting that ls26/ls27 belong to a novel protein family. Programmes were also tested which were designed to identify conserved protein motifs (for example Prosite), but again there was no clue as to the kind of activity that the ls26/ls27 proteins may have. Thus, based on the lack of sequence homology, the ls26/ls27
30 proteins do not appear to have any known catalytic activity (protein kinase, phosphatase,...).

- 10 -

Example 2

To investigate the function of the ls26/ls27 proteins, the two cDNAs were cloned in the bacterial expression vector pMalc2 downstream of the malE gene. The fusion proteins malE-ls26 and malE-ls27 were expressed in and purified from *E. coli*. Since ls26 and ls27 are very similar in sequence and probably correspond to pseudoalleles, which are quite common in *Xenopus*, one concentrated on the characterization of ls26 and then confirmed the results obtained using ls27. It was found that the injection into *Xenopus* oocytes of small amounts of recombinant malE-ls26 protein (10 ng) was capable of inducing oocyte maturation considerably faster than the same amount of injected malE-Mos or than progesterone treatment (Fig. 2). This experiment using the fusion protein confirmed the results observed with mRNA in vitro transcribed from the ls26 cDNA clone regarding the potency of this novel protein to induce cell cycle activation in *Xenopus* oocytes. It was also found that injection of only 0.5 ng of malE-ls26 per oocyte was still capable of inducing oocyte maturation. The availability of purified malE-ls26 protein also allowed to test the capability of ls26 to induce oocyte maturation in the presence of protein synthesis inhibitors. Preincubation of the oocytes with cycloheximide totally blocked progesterone-induced maturation, consistent with the known essential requirement for translation of maternal mRNAs stored in the oocytes for progesterone to induce maturation. In the same experiment, cycloheximide preincubation also blocked malE-Mos-induced GVBD but it had no effect on ls26-triggered oocyte maturation (Fig. 2).

Example 3

To further characterize the activity of the ls26 protein, the kinetics of activation of MAP kinase (MAPK) and cdc2/cyclin B (MPF) in oocytes induced to mature by malE-ls26 were investigated. MAPK and MPF are normally activated during oocyte maturation and their activation can be

- 11 -

detected in oocyte lysates either by immunoblot with anti-MAPK and anti-cdc2 antibodies or by direct measurement of the in vitro kinase activity using MBP and histone H1 as substrates for MAPK and MPF, respectively (Fig. 3). As expected from previous work, we observed that progesterone treatment activates both MAPK and MPF at about the same time, whereas malE-Mos injection activates MAPK well before MPF activation. Interestingly, injection of malE-Is26 rapidly activates MPF somewhat before MAPK. Moreover, MPF appears to be transiently activated by Is26, but the significance of this observation is unclear. In cycloheximide-treated oocytes, the Is26-induced activation of MAPK is very much reduced whereas the activation of MPF is apparently unaffected. This result indicates that the effect of Is26 is more related to MPF activation than to MAPK activation. As expected, cycloheximide totally blocked progesterone-induced activation of both MPF and MAPK, whereas in the case of Mos only MPF but not MAPK activation was compromised by cycloheximide.

Example 4

The observation that Is26 can consistently induce oocyte maturation and the activation of MPF independently of new protein synthesis is quite remarkable as only proteins that act very late in the activation pathways, such as cyclins (cdc2 binding and activating subunits) or direct cdc2/cyclinB activators such as the cdc25 phosphatase have been shown to have this strong effect. In order to address whether Is26 can directly associate with and/or modify the activity of cdc2/cyclin B complexes, pull-down experiments were performed. For this purpose, extracts prepared from insect cells infected with cdc2-expressing baculovirus were incubated with either malE-Is26 bound to amylose beads or the equivalent amount of cyclin B bound to nickel beads. After extensive washing, the proteins that remained bound to the beads were analyzed by immunoblot using anti-cdc2 antibodies and in vitro histone H1 kinase assay. We found that Is26 bound to cdc2 with almost the same efficiency of cyclin B (Fig. 4). When the

- 12 -

cyclin B pull-down was done in the presence of a 2-fold molar excess of soluble malE-Is26, we observed a reduction in the amount of cdc2 bound to cyclin B which also correlated with the expected decrease in the kinase activity of the complexes. This suggests that Is26 may compete with cyclin B to bind to cdc2. By immunoblot using anti-malE antibodies we confirmed that cyclin B was capable of binding Is26 in the presence of cdc2. Our results indicate that Is26 can strongly bind to cdc2 and probably also to cyclin B, but we do not know whether Is26 can complex to cdc2/cyclin B.

10 The interaction between Is26 and cdc2 was confirmed using ³⁵S-methionine-labelled cdc2 prepared by coupled transcription/translation in rabbit reticulocyte lysates. We also confirmed in pull-down experiments with rabbit reticulocyte lysates that Is26 can directly bind to B-type cyclins (Fig. 5).